PCT ORLD INTELLECTUAL PROPERTY ORGANIZATIO
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C07K 13/00, C12N 9/20, 9/16
C12N 15/55, A61K 37/54

A1

(11) International Publication Number: **WO 91/18923**

(43) International Publication Date: 12 December 1991 (12.12.91)

(21) International Application Number: PCT/SE91/00381

(22) International Filing Date: 30 May 1991 (30.05.91)

(30) Priority data:
9001985-2 1 June 1990 (01.06.90) SE

(71) Applicant: AKTIEBOLAGET ASTRA [SE/SE]; S-151 85 Södertälje (SE).

(72) Inventors: BJURSELL, Gunnar ; Riddar Tures v. 13, S-433 31 Partille (SE). BLÄCKBERG, Lars ; Scharinsvägen 18, S-902 44 Umeå (SE). CARLSSON, Peter ; Kommendörsg. 30 A, S-414 59 Göteborg (SE). ENERBÄCK, Sven ; Slätthultsliden 13 A, S-431 69 Mölndal (SE). HERNELL, Olle ; Gimoborgsv. 14, S-902 40 Umeå (SE). NILSSON, Jeanette ; Sveag. 16, S-413 14 Göteborg (SE). OLIVECRONA, Thomas ; Kungsg. 21 A, S-902 45 Umeå (SE).

(74) Agents: MIKSCHE, Gerhard et al.; AB Astra, Patent Department, S-151 85 Södertälje (SE).

(81) Designated States: AT, AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH, CH (European patent), CI (OAPI patent), CM (OAPI patent), DE, DE (European patent), DK, DK (European patent), ES, ES (European patent), FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), GN (OAPI patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU, LU (European patent), MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL, NL (European patent), NO, PL, RO, SD, SE, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent).

Published
With international search report.

(54) Title: DERIVATIVES OF HUMAN BILE-SALT STIMULATED LIPASE, AND PHARMACEUTICAL COMPOSITIONS CONTAINING THEM

(57) Abstract

The present invention relates to new DNA sequences, to new proteins coded for by such DNA sequences, and to the use as further described below of such proteins. The invention also encompasses vectors, such as plasmid constructs, comprising such DNA sequences, being capable of expressing the desired enzyme. The invention also includes host organisms transfected with such constructs e.g. bacteria yeast, mammalian cells, and transgenic animals. The invention also includes processes for the preparation of the novel products of the invention. The new proteins of the invention are related to an enzyme, known i.a. as human bile salt-stimulated lipase.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

DERIVATIVES OF HUMAN BILE-SALT STIMULATED LIPASE, AND PHARMACEUTICAL COMPOSITIONS CONTAINING THEM

Field of the invention

The present invention relates to new DNA sequences, to new proteins coded for by such DNA sequences, and to the use as further described below of such proteins. The invention also encompasses vectors, such as plasmid constructs, comprising such DNA sequences, being capable of expressing the desired enzyme. The invention also includes host organisms transfected with such constructs e.g. bacteria yeast, mammalian cells, and transgenic animals. The invention also includes processes for the preparation of the novel products of the invention. The new proteins of the invention are related to an enzyme, known i.a. as human bile salt-stimulated lipase.

Background to the invention

The human lactating mammary gland synthesizes and secretes with the milk a bile salt-stimulated lipase (BSSL) [1] that, after specific activation by primary bile salts [2, 57, 58], contributes to the breast-fed infant's endogenous capacity of intestinal fat digestion [3-5]. This enzyme, which accounts for approximately 1% of total milk protein [6], is a non specific lipase; in vitro it hydrolyses not only tri-, di- and monoacylglycerols, but also cholesteryl-, and retinyl esters, and lysophosphatidylglycerols [7-10]. Furthermore, its activity is not restricted to emulsified substrates, but micellar and soluble substrates are hydrolyzed at similar rates [11].

BSSL is not degraded during passage with the milk through the stomach, and in duodenal contents it is protected by bile salts from inactivation by pancreatic proteases such as trypsin and chymotrypsin [2,11]. It is, however, inactivated when the milk is pasteurized, e.g. heated to 62.5 °C, 30 min [12]. Model experiments in vitro suggest that the end products of triacylglycerol digestion are different in the presence of BSSL [5,7]. Due to lower intraluminal bile salt concentrations during the neonatal period [13,14] this may be beneficial to product absorption [5,15].

The carboxylic ester hydrolase (CEH) of human pancreatic juice [16] seems functionally to be identical, or at least very similar, to BSSL [8]. They also share common epitopes [8,17], have identical N-terminal amino acid sequences [17] and are inhibited by inhibitors of serine esterases, e.g. eserine and diisopropylfluorophopsphate [6,8,16]. It has been hypothesized that the two enzymes are products of the same gene [18,19]. The observed molocular size difference [8,19] could be explained by different patterns of glycosylation, as recently suggested [17].

Dietary lipids is an important source of energy. The energy-rich triacylglycerols constitute more than 95% of these lipids. Some of the lipids, e.g. certain fatty acids and the fat soluble vitamins, are essential dieatary constituents. Before gastro-intestinal absorption the triacylglycerols as well as the minor components, i.e. esterified fat-soluble vitamins and cholesterol, and diacylphosphatidylglycerols, require hydrolysis of the ester bonds to give rise to less hydrophobic, absorbable products. These reactions are catalyzed by a specific group of enzymes called lipases.
In the human adult the essential lipases involved are considered to be gastric lipase, pancreatic colipase-dependent lipase (tri- and diacylglycerol hydrolysis), pancreatic phospholipase A2 (diacylphosphatidylglycerols) and carboxylic ester hydrolase (cholesteryl- and fat soluble vitamin esters). In the breast-fed newborn bile salt-stimulated lipase plays an essential part in the hydrolysis of several of the above mentioned lipids. Together with bile salts the products of lipid digestion form mixed micelles from which absorption occurrs (3-5).

Common causes of lipid malabsorption, and hence malnutrition, are reduced intraluminal leavels of pancreatic colipase-dependent lipase an/or bile salts. Typical examples of such lipase deficiency are patients suffering from cystic fibrosis, a common genetic disorder resulting in a life-long deficiency in som 80% of the patients, and chronic pancreatitis, often due to chronic alcoholism.

The pancreatic and liver functions are not fully developed at birth, most notably in infants born before term. Fat malabsorption, for physiological reasons, is a common finding and thought to result from low intraluminal pancreatic colipase-dependent lipase and bile salt concentrations (3,4,13-15). However, because of BSSL, such

malabsorption is much less frequent in breast-fed infants than in infants fed pasteurized human milk of infant formulas (3-5, 12, 59, 60, 61). This is one reason why it has been advocated that newborn infants, particularly preterm infants, that cannot be fed their own mothers milk should be fed non-pasteurized milk from other mothers (12).

The present treatment of patients suffering from a deficiency of pancreatic lipase is the oral administration of very large doses of a crude preparation of porcine pancreatic enzymes. Colipase-dependent pancreatic lipase is inactivated by low pH. Such conditions are prevalent in the stomach, with the result that orally administered pancreatic lipase is virtually completely inactivated on the passage through the stomach to the gut. Therefore, this effect cannot be completely overcome by the use of large doses of enzyme. The large doses administered are inadequate for most patients, and the preparations are impure and unpalatable. Certain tablets have been formulated which pass through the acid regions of the stomach and discharge the enzyme only in the relatively alkaline environment of the jejunum. However, many patients suffering from pancreatic disorders have an abnormally acid jejunum and such tablets may fail to discharge the enzyme and may therefore be ineffective. Moreover, since the preparations presently on the market are of a non-human source there is a risk of immunoreactions that may cause harmful effects to the patients or result in reduced therapy efficiency.
A further drawback with the present preparations is that their content of other lipolytic activities than colipase-dependent lipase are not stated. In fact, most of them contain very low levels of CEH/BSSL-activity. This may be one reason why many patients, suffering from cystic fibrosis in spite of supplementation therapy, suffer from deficiencies of fat soluble vitamins and essential fatty

acids.

Thus, there is a great need for products with properties and structure derived from human lipases and with a broad substrate specificity, which products may be orally administered to patients suffering from deficiency of one or several of the pancreatic lipolytic enzymes. The products of the present invention fulfil this need by themselves or in combination with other lipases or in combination with preparations containing other lipases. Furthermore, for some human infants there is an obvious need to improve fat utilization from conventional infant formulas, or pasteurized human milk from so-called milk banks.

BSSL has several unique properties that makes it ideally suited for substitution and supplementation therapy:

It has been designed by nature for oral administration. Thus, it resists passage through the stomach and is activated in contents of the small intestine.

Its specific activation mechanism should prevent hazardous lipolysis of food or tissue lipids during storage and passage to its site of action.

Due to its broad substrate specificity it has the potential to, on its own, mediate complete digestion of most dietary lipids, including the fat soluble vitamin esters.

BSSL may be superior to pancreatic colipase-dependent lipase to hydrolyze ester bonds containing long-chain polyunsaturated fatty acids.

In the presence of gastric lipase and in the absence of, or at low levels of colipase-dependent lipase BSSL can ascertain a complete triacylglycerol digestion in vitro even if the bile salt levels are low such as in newborn infants. In the presence of BSSL the end products of triacylglycerol digestion become free fatty acids and free glycerol rather than free fatty acids and monoacylglycerol generated by the other two lipases (5). This may favour product absorption particularly when the intraluminal bile

salt levels are low (3,15).
From a historical point of view infant formulas have been developed, and improved, from the concept that their composition should be as similar to that of human milk as possible. It is desirable to supplement such formulas.

The utilization for supplementation, substitation or therapy of bile salt-stimulated lipases (BSSL), or of proteins with the essential functions of BSSL, requires however access to quantities of the product on a large technical scale. It is not possible in factory scale to rely on natural sources such as milk as starting material. Besides the problem mentioned above with inactivation of BSSL during pasteurization, there is the additional risk of contamination of material from a natural source with infectious agents, e.g. vira such as HIV virus and CMV. There is, accordingly, a need for large scale access to products having BSSL properties. The present invention provides such products and methods for their preparation.

Prior art references are given later in this specification.

The invention

The present invention is based on the cloning of cDNA coding for BSSL derived from human mammary gland. We have also isolated, from human pancreas, a partial cDNA coding for CEH. Deduced amino acid sequences from the human cDNA's and comparison with CEH from other species, support the interpretation that BSSL and CEH are identical.

As will be further detailed below, it was surprisingly found that the structure of the protein as deduced from the cDNA sequence is quite different from the structure of other lipases. The structure proved unexpectedly to be more like the structure of typical esterases, such as cholinesterase.

With reference to Figure II and Figure VII, products of the invention are:

a) a protein as defined by the amino acid sequence 1-722 in Fig. VII,

b) a protein as defined by the amino acid sequence 1-535 in Fig. VII,

c) a protein as defined by the amino acid sequence 1-278 in Fig. VII,

d) a protein as defined by the amino acid sequence 1-341 in Fig. VII,

e) a protein as defined by the amino acid sequence 1-409 in Fig. VII,

f) a protein as defined by the amino acid sequence 1-474 in Fig. VII,

g) combinations of proteins defined under b) - f) e.g. as defined by the amino acid sequence in positions 1-278, 279-341, 279-409, 279-474, 342-409, 342-474 and 536-722.

h) combinations of proteins defined under b) - g) in combination with one or more of the repeats according to Figure V,

i) a protein as defined under a) - h) possessing an additional, N-terminal amino acid, namely methionine, and functionally equivalent variants and mutants of the proteins defined in a) - i) above;

It should be noted that the proteins under a, b, c, d, e, f, h and i above will not be identical in all respects to naturally occurring BSSL, but they will exhibit one or

more of the critical functions of naturally occurring BSSL. Critical functions are given below.

j) a DNA sequence coding for the proteins defined in a, b, c, d, e, f, h and i above.

k) a DNA sequence according to Fig. II, defined by the following nucleotide numbers in Fig. II:

a) a DNA sequence 151-2316 according to Fig. II, coding for the protein defined by the amino acid sequence 1-722 in Fig. VII,
b) a DNA sequence 151-1755 according to Fig. II, coding for the protein defined by the amino acid sequence 1-535 in Fig. VII,
c) a DNA sequence 151-985 according to Fig. II, coding for the protein defined by the amino acid sequence 1-278 in Fig. VII,
d) a DNA sequence 151-1172 according to Fig. II, coding for the protein defined by the amino acid sequence 1-341 in Fig. VII,
e) a DNA sequence 151-1376 according to Fig. II, coding for the protein defined by the amino acid sequence 1-409 in Fig. VII,
f) a DNA sequence 151-1574 according to Fig. II, coding for the protein defined by the amino acid sequence 1-474 in Fig. VII,
g) a DNA sequence 986-1172 according to Fig. II, coding for the protein defined by the amino acid sequence 279-341 in Fig. VII,
h) a DNA sequence 986-1376 according to Fig. II, coding for the protein defined by the amino acid sequence 279-409 in Fig. VII,
i) a DNA sequence 986-1574 according to Fig. II, coding for the protein defined by the amino acid sequence 279-474 in Fig. VII,

j) a DNA sequence 1173-1376 according to Fig. II, coding for the protein defined by the amino acid sequence 342-409 in Fig. VII,

k) a DNA sequence 1173-1574 according to Fig. II, coding for the protein defined by the amino acid sequence 342-474 in Fig. VII.

Significant functions of proteins of the invention are

a) suitable for oral administration,

b) being activated by specific bile salts,

c) acting as a non-specific lipase in the contents of the small intestines, that is being able to hydrolyze lipids relatively independent of their chemical structure and physical state (emulsified, micellar, soluble).

d) Ability to hydrolyze triacylglycerols with fatty acids of different chain-length and different degreee of unsaturation.

e) Ability to hydrolyze also diacylglycerol, monoacylglycerol, cholesteryl esters, lysophosphatidylacylglycerol, and retinyl and other fat soluble vitamin-esters.

f) Ability to hydrolyze not only the sn-1(3) ester bonds in a triacylglycerol but also the sn-2 ester bond.

g) Ability to interact with not only primary but also secondary bile salts.

h) Dependency on bile salts for optimal activity.

i) Stability so that gastric contents will not affect the catalytical efficiency to any substantial degree.

j) Stability towards inactivation by pancreatic proteases, e.g. trypsin, provided bile salts are present.

k) Ability to bind to heparin and heparin derivatives, e.g. heparan sulphate.

l) Ability to bind to lipid-water interphases.

m) Stability to permit lyophilization.

n) Stability when mixed with food constituents such as in human milk, or milk formula.

The critical functions for supplementation, substitution, or therapy are these according to a), c), d), e), f), i), j) and l). For other purposes, not all critical functions may be necessary.

For expression of the proteins indicated above, the appropriate DNA sequence indicated above will be inserted into a suitable vector which then is introduced into a suitable host organism. The said vector will also have to comprise appropriate signal and other sequences enabling the organism to express the desired protein.

Suitable expression organisms:

With the recombinant DNA techniques it is possible to clone and express a protein of interest in a variety of prokaryotic and eukaryotic host organisms. Possible expression organisms are bacteria, simple eukaryots (yeast), animal cell cultures, insect cell cultures, plant cell cultures, plants and transgenic animals. Each individual system has its own particular advantages and disadvantages. The simple conclusion is that every gene to be expressed is a unique problem and no standard solution

is available.

Commonly used bacterial systems are E. Coli, Bacillus subtilis, Streptomyces. Commonly used yeasts are saccharomyces, and Pichia Pastoris. Commonly used animal cells are CHO cells and COS cells. Commonly used insect cell cultures are Drosophila derived cells.

Commonly used plant is the tobacco plant. Possible transgenic animals are goat and cow.

Possible bacterial vectors are exemplified by pUC and protein A-vectors.
Possible yeast vector is exemplified by pMA 91.
Possible insect vectors are devired from Baculo-virus.
Possible animal cell vectors are derived from SV/40.
Possible plant vectors are derived from the Ti-plasmid.
In every system, both natural and synthetic promoters and terminators can be used.

It is understood that depending on the choice of expression system, the expressed protein may contain an additional N-terminal amino acid (methionine), contain a few extra amino acids, or be fused to a heterologous protein, (e.g. protein A), or differ from the naturally occurring protein with respect to glycosylation. Furthermore, the vectors may also contain signal sequences in order to export the protein to the periplasm or to the culture medium.

Thus, further aspects of the invention are:

a) a vector comprising a DNA sequence coding for a protein as specified above,

b) a host organism comprising a DNA sequence as specified above,

c) a process for the production of a protein as specified above, by growing a host organism containing a vector as specified under a) above and isolating the protein.

Methods from purification are based on the expression system used (e.g. protein A/IgG) and/or on methods used for purification of the naturally occurring enzyme, as described in reference 6.

Additional aspects of the invention are:

- a pharmaceutical composition comprising a protein as specified above,

- the use of a protein as specified above for the manufacture of a medicament for the treatment of a pathological condition related to exocrine pancreatic insufficiency,

- the use of a protein as specified above for the manufacture of a medicament for the treatment of cystic fibrosis,

- the use of a protein as specified above as a supplement to an infant food formulation,

-the use of a protein as specified above for the manufacture of a medicament for the treatment of chronic pancreatitis,

-the use of a protein as specified above for the manufacture of a medicament for the treatment of fat malabsorption of any etiology,

-the use of a protein as specified above for the manufacture of a medicament for the treatment of malabsoption of fat soluble vitamins,

-the use of a protein as specified above for the manufacture of a medicament for the treatment of fat malabsorption due to physiological reasons, e.g. in new-

born infants.

The DNA sequence in Fig. II from position 151 up to and including position 2316 is the sequence coding for the entire protein., The sequence from position 2317 up to and including position 2415 is not translated to protein, but is included in exon d identified in Table 2 below.

In one embodiment of the invention, the protein as defined in pharagraphs a) - i) above is provided in isolated form and/or in substantially pure form.

The DNA sequences as defined in paragraphs a) - k) above are in one embodiment of the invention provided in isolated form and/or in substantially pure form

Experimental part

Abbreviations

aa, amino acid; bp, base pair; BSSL, bile salt-stimulated lipase; c-AMP, cyclic adenosine monophosphate; CEH, carboxylic ester hydrolase; Da, dalton; c ^{7}GTP, 7-deaza-2-deoxyguanosine 5'triphosphate; EDTA, ethylene diamine tetraacetate; kb, kilobases; MOPS, 3-N-morpholino-propanesulfonic acid; nt, nucleotide; PAGE, polyacrylamide gel electrophoresis; SDS, sodium dodecyl sulfate; SSC, NaCl citrate, xGal, 5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside.

Enzymes

Bile salt-stimulated lipase EC 3.1.1.3
Carboxylic ester hydrolase EC 3.1.1.1

Material and Methods

A. Enzyme and antibody preparation

BSSL was purified from human milk as previously described [6]. When used for antibody production the enzyme was further purified by SDS-PAGE. The protein band corresponding to the lipase was, after staining with Coomassie Brilliant blue, electroeluted from the gel. Twentyfive µg of purified enzyme, together with an equal volume of Freund's complete adjuvant, was used for a first i.c. injection and the same amount of enzyme with incomplete adjuvant for the subsequent monthly booster injections. The rabbits were bled about two weeks after each booster and sera prepared and stored at -20 °C.

B. Preparation of tryptic fragments and amino acid sequence analysis

Three mg of purified BSSL was dissolved in 1 ml of 0.1M Tris-Cl buffer, pH 8.5, containing 6M guanidinium hydrochloride and 2 mM EDTA. Dithioerythritol was added to 5mM. After incubation at 37 °C for 2h, 300 µl 50 mM iodoacetate was added. After 90 min incubation at 25 °C in darkness the reduced and carboxymethylated enzyme was desalted on a Sephades G-25 column, equilibrated with 0.5M ammonium bicarbonate. Thirty µg of tosyl-L-phenylalanine chloromethane treated bovine trypsin (Worthington diagnostics system Inc., Freehold, NJ, USA) was added before lyophilization. The lyophilized protein was dissolved in 4 ml 0.1M ammonium bicarbonate and an additional 90 µg of trypsin was added. After 5h incubation at 37 °C the protein was again lyophilized. The tryptic digest was dissolved in 0.1% trifluoroacetic acid (2mg/ml). Three hundred µg of trypsinated BSSL was chromatographed on HPLC using a C-18 reversed phase column and eluted with a gradient of 0-50% acetonitrile in 0.1%

trifluoroacetic acid. Peptide collection was monitored by continuous recording of the absorbance at 215 nm. Peptides to be sequenced were further purified by rechromatography using the same column with adjusted gradients.

Samples of peptide fragments to be sequenced were dried under nitrogen to remove acetonitrile and applied to the sequencer. For N-terminal sequence analysis, native BSSL was dissolved in 0.1% acetic acid. Sequence analyzes were performed on an Applied Biosystems Inc. 477A pulsed liquid-phase sequencer and on-line PTH 120A analyzer with regular cycle programs and chemicals from the manufacturer. Initial and repetitive yields, calculated from a sequenced standard protein, ß-lactoglobulin, were 47 and 97%, respectively.

C. Isolation of RNA

Samples of human pancreatic adipose and lactating mammary gland tissues were obtained at surgery and immediately put into guanidinium thiocyanate (1-5 g in 50 ml). Total RNA was extracted as described by Chirgwin [20]. Poly(A)-RNA was prepared by chromatography on oligo-deoxythymidilate-(oligo(DT))-cellulose column [21].

D. Construction and screening of cDNA libraries

Approximately 15 µg poly-adenylated RNA from human pancreas was denatured with methyl mercuric hydroxide [22] and primed with oligo (dT) $_{12-18}$ primers (Pharmacia, Uppsala, Sweden), and reversely transcribed using standard procedures [23]. Second-strand synthesis was carried out according to Gubler and Hoffman [24], except that DNA ligase and ß-NAD were omitted, and the reaction temperature was set at 15 °C. Excess RNA was digested with RNAse A (50µg/ml), and the double-stranded cDNA was treated with EcoRI methylase [25]. Ends were blunted with Klenow enzyme. After ligation to EcoRI linkers and

cleavage with EcoRI the cDNA was fractionated on a Sepharose 4B-Cl column. The void volume fraction was precipitated with ethanol and the cDNA ligated into the EcoRI site of a phosphatase treated gtll vector [26]. In vitro packing yielded more than $7x10^5$ recombinants.

A cDNA library from human mammary gland, derived from tissue obtained from a women at the eighth month of pregnancy, was purchased (Clontech Laboratories, Inc., Palo Alto, CA; USA).

Phages from the cDNA libraries were plated at $5x10^4$ plaque forming units per 120-mm dish. The antiserum was diluted to a ratio of 1:3200 and screening was performed according to Young and Davis [27]. Alkaline-phosphatase-conjugated goat-anti-rabbit antibodies were used as second antibodies (Bio-Rad, Richmond, CA USA). To isolate clones corresponding to the 5'-end of the mRNA, nucleic acid hybridization was done under standard conditions [23] using a subcloned fragment from one of the immunopositive clones as a probe.

E. RNA analysis

Electrophoresis was carried out in a 1% agarose gel in 40mM MOPS buffer pH 7.0 after denaturation with glyoxal and dimethylsulfoxide [28]. Glyoxalated total RNA was then transferred to nitrocellulose filters [29]. The blots were probed with subclones of BSSL and CEH recombinants that were labelled by the oligo-labeling technique [30]. Prehybridization and hybridization were carried out with 50% formamid at 46 °C [23]. Posthybridization washes were performed at high stringency (0.1% SDS and 0.1xSSC at 60° C). (1xSSC, 0,15M NaCl, 0.0015M Na_3 citrate, pH 7.6).

F. Nucleotide sequence

cDNA inserts from BSSL and CEH recombinants were either directly cloned into M3mp18 and mp19 after sonication and size fractionation or some of them were further subcloned into pTZ19R after digestions with PstI, BstXI, NarI, SmaI and AhaII. The nucleotide sequence was determined by the dideoxy chain-termination method [31]. The GC-rich repeats (see below) were also sequenced with TaqI polymerase and dc ^{7}GTP. Both strands were sequenced. Sequence information was retrived from autoradiograms by use of the software MS-EdSeq as described by Sjöberg et al [55].

G. Amino acid sequence predictions and homologies.

To predict the corresponding amino acid sequence of the cDNA inserts, codon usage of different reading frames was compared according to Staden and gave one open reading frame [32]. Homologies were searched for with the programs of the UWGCG software package [33].

Results and Discussion

A. Sequences of tryptic fragments and the N-terminus of BSSL

Trypsin digestion of purified BSSL resulted in approximately 50 fragments as judged by the number of peaks obtained during the HPLC-chromatography (Fig. I). The peaks were collected and the indicated peaks which could be isolated in a highly purified state and in reasonable quantities, were sequenced. The resulting sequences are shown in Table I. In addition the 30 most N-terminal residues were sequenced (Fig. II), and they confirm the previously reported sequence of Abouakil et al. (30 residues) [17].

The 22 residues long N-terminal sequence reported from Wang & Johnson [34] is a glycine in our report and a

lysine in Wang & Johnson's.

B. Nucleotide sequence of BSSL

For construction of the λgt11 cDNA library we used polyadenylated RNA from human pancreas. Initially four immunopositive clones were isolated, and then this pancreatic expression cDNA library was screened with antiserum against BSSL. Nucleotide sequence analysis of the four clones showed that they are in perfect agreement and correspond to the 3'-end of the mRNA. They all begin with a poly A tail and differ only in length; the longest insert, designated ACEH, spans 996 bp.

A cDNA library from human mammary gland was screened with antiserum, and the pancreas clone ACEH as probe. Positive clones were isolated from both screenings, which all originate from the 3'-end. The longest mammary gland clone, designated ABSSL, reaches 2100 bp upstream. It contains four of the sequenced tryptic fragments (Fig. II), but do not include the N-terminal amino acid sequence. To extend the sequence beyond the translation start, the mammary gland cDNA library was rescreened with a 118 bases long probe derived from the most 5'proximal part of ABSSL. One clone was isolated that continued a further 328 nucleotides upstream. It matched the N-terminal amino acid sequence, and contained the remaining tryptic fragment. As shown in Fig. II, the cDNA is 2428 nucleotides long and contains 81 bases upstream from the first ATG codon. The polyadenylation signal, AATAAA is located 13 nucleotides upstream from the poly A tail and the termination codon TAG was found at nucleotide 2317 followed by a 3'-untranslated region of 112 bp.

A GC rich region consisting of 16 repeats of 33 bases was found in the 3'-end of the sequence between base 1756 and 2283. The nucleotide sequence of the repetition, shown in Fig. III, consists of six identical repetitions surrounded by ten repetitions with different number of

substitutions that have probably occurred after several duplications. The low number of substitutions suggests that these repetitions have appeared late during evolution.

C. Tissue distribution of expression

RNA from human lactating mammary gland, pancreas, adipose tissue and from a human hepatoma celline (HepG2) was analyzed by Northern blotting. The size of the messenger was determined to be approximately 2.5 kb in both lactating mammary gland and pancreas. No signal could be detected in the lanes with RNA extracted from HepG2 or adipose tissue (Fig. IV).

Since the mRNA used for the mammary gland library was obtained from a female in her 8th month of pregnancy, it is evident that transcription and probably translation of the BSSL gene is turned on before partus, in agreement with previous findings on BSSL secretion before partus [35]. See Figure IV.

D. Amino acid sequence of BSSL

Assessed by SDS-PAGE the molecular mass has been reported to be 107-125 kDa [8,36] and by analytical ultracentrifugation to be 105 kDa [37]. The enzyme, as deduced from the cDNA, consists of 722 amino acid residues (Fig. II) which, giving a molecular mass of 76.271 Da, indicates that the enzyme contains at least 15-20% carbohydrate. The leader sequence is 23 residues long. A tentative active site serine residue is localized to serine-217 (Fig. V). The sequence around this serine accord with the consensus active site sequence of serine-hydrolases [38]. It has recently been it was proposed that basic residues found close to the active site serine may be involved in the cleavage of esterbonds in acylglycerols by lipases [39]. It is interesting to note that such

residues are not present in BSSL. The single tentative N-glycosylation site is localized only seven residues from the serine. The degree of glycosylation [6,16] suggests that the enzyme contains 0-linked carbolydrate. There are numerous sites where such glycosylation could have occurred. The amino acid composition based on purified enzyme has shown a high content of proline residues [6]. The amino acid sequence obtained from cDNA confirms this. Moreover, most of the proline residues are localized in the 16 repeats of 11 residues each, constituting the main part of the C-terminal half of the enzyme.

E. Comparison of the enzymes in mammary gland (BSSL) and pancreas (CEH)

BSSL of human milk and human pancreas CEH have previously been shown to be similar, if not identical. The present data strongly suggests that the two enzymes are products of the same gene. The nucleotide sequence of the cDNA clones shows that the pancreatic clone ACEH is identical with the mammary gland clone ABSSL from the poly A tail and 996 bases towards the 5'-end, including the sequence coding for the proline rich repeats. Northen blot gave a single band of 2.5 kb in RNA from pancreas and lactating mammary gland (Fig. IV). Genomic Southern blots further support the idea that only one gene codes for BSSL and CEH. The difference in mobility on SDS-PAGE between BSSL and CEH can be explained as a consequence of different glycosylation or differential splicing.

The similarity of BSSL to the rat and bovine enzymes (see below) and to results from genomics blots support the possibility that differential splicing cannot account for the mobility difference. Since the C-terminal sequence has not been confirmed on the protein level there is a less likely possibility that CEH may be processed by a proteolytic clearage in the C-terminal end.

So far as we know pancreatic enzymes that obviously correspond to CEH have often been named after species and the particular substrates used to determine their respective activities; lysophospholipase, cholesteryl esterase, sterol ester hydrolase, non-specific lipase, carboxyl ester lipase and cholesteryl ester hydrolase. Available data are compatible with the view that all these activities described originates in one and the same functional entity [42,43]. This illustrates the broad substrate specificity and the relevance of designating them as non-specific lipases. When the sequence of human BSSL/CEH is compared to the sequence of lysophospholipase from fat pancreas [40] and cholesterol esterase from bovine pancreas [41] extensive similarities are found that extend about 530 residues from the N-terminal (Fig V); but they differ in the part of the molecule where the repeats occur. The rat enzyme has only four repeats and the bovine three. Hence the human enzyme is a considerable longer peptide.

Moreover, striking similarities were found between BSSL and a number of typical esterases, e.g. acetyl choline esterases from several species, including man and Drosophila, and carboxyl esterases (Fig. VI). These similarities were restricted to the N-terminal 300 residues of BSSL which includes the tentative active site serine-residue. A similarity to acetyl choline esterase has been predicted from the fact that BSSL is inhibited by typical choline esterase inhibitors [6, 8, 16]. With the possible exception of the rat liver carboxyl esterase [45), none of these similar enzymes has been shown to have the same bile-salt dependency as BSSL; this suggests that the structural basis for this property resides in the C-terminal part of the protein. Moreover, BSSL can efficiently attack emulsified substrates which is not a known characteristic of the similar esterases. For this activity bile salt is a prerequisite.

The predicted sequence for human BSSL was compared with other well characterized mammalian lipases. Apart from the consensus sequence around the active site serine (G-X-S-X-G), no obvious similarities were found [44].

In addition to the similarities with other enzymes, there also significant similarities to one c-AMP dependent protein from Dictyostelium discoideum [46] as well as to thyroglobulin from several species (Fig VI) [47-49]. The similarities between BSSL and thyroglobulin, which comprise the active site region but not the active site itself, indicate that these highly concerved stretches of amino acids are of more generalized importance than merely supporting the enzymatic activity of esterases.

In conclusion, human milk BSSL consists of 722 amino acid residues. Available data strongly indicate that its peptide chain is identical to that of pancreatic CEH, and they are coded for by the same gene. The strongest evidence is that the nucleotide sequences of their 3 - ends and their N-terminal amino acid sequences are identical. The striking homologies found to rat pancreatic lysophospholipase and bovine pancreatic cholesterol esterase support the hypothesis that also these enzymes are functionally identical. However, as it has been suggested, the different molecular sizes found among species are not due to differences merely in glycosylation; instead they reflect a variable number of an eleven amino acid repeat. The similarity of the active site sequence between these esterases suggests that these proteins derive from a common ancestral gene.

With reference to Figures I-VII, the following legends are given.

Figure I: Separation of the tryptic digest of BSSL on HPLC

Purified BSSL was treated with trypsin and chromatographed on HPLC as described in Materials and Methods. The indicated peaks were collected and purified further by a rechromatograph and their amino acid sequence determined.

Figure II: The cDNA nucleotide sequence and the deduced amino acid sequence for human bile salt-stimulated lipase:

The cDNA is 2428 bases long. The N-terminal 23-codon sequence (nt82-150) starting with an ATG, is interpreted as a leader peptide since the N-terminal amino acid sequence of the mature protein starts at codon 24 (nt 151, Ala). The leader peptide is underlined. The sign * indicates the starting point of an exon. The sign ‡ indicates the starting point of the repetition part.

Figure III: The nucleotide sequence of the C-terminal GC-rich repetitions in the bile salt-stimulated lipase:

Substitutions are indicated by a *.

Figure IV: Northern blot hybridization

Northern blot analysis of total RNA isolated from human lactating mammary gland, pancreas, adipose tissue and a human hepatoma cell line (HepG2). Total TNA (10µg) from lactating mammary gland (lane A), pancreas (lane B), adipose tissue (lane C) and HepG2 (lane D) were electrophoresed in a 1% agarose gel in 40mM MOPS buffer at Ph 7.0 after denaturation of RNA in 1M glyoxal, 50% dimethylsulfoxide and 40mM MOPS. The glyoxalated RNA was then transferred to nitrocellulose paper for hybridization with [^{32} P] labeled BSSL cDNA (ABSSL).

Figure V: Comparison of the deduced amino acid sequence from human milk BSSL, rat pancreatic lysophospholipase (Ratlpl) [40] and bovine pancreatic cholesterol esterase (Bovceh) [41]:

The serine residues involved in the active site are indicated by a *, and the # indicates the single possible N-glycosylation signal of the protein. The direct repeats of amino acid sequences are boxed. Matching sequences are denoted in capital letters, matching sequences between two enzymes are denoted in small letters and mismatching with a dot.

Figure VI: Comparison of the primary structure of BSSL to other esterases, thyroglobuline and to one c-AMP dependent enzyme from Dictyostelium discoideum:

BSSL: bile salt stimulated lipase from human, Cheshum: cholinesterase from fetal human tissue [50], Torpace: acetylcholinesterase from Torpedo marmorata [51], Drosceh: carboxylic ester hydrolase from Drosophila melaogaster [52], Ratlivce: carboxyl esterase from rat liver [53], Drosace: acetylcholinesterase from Drosophila melaogaster [54], Thyrhum: thyroglobulin from human [49] and Dict.Di: c-AMP dependent enzyme from Dictyostelium discoideum [46]. There are 7 different domains that show similarities between the enzymes. Boxes enclose residues which are identical and small letters in the consensus sequence indicate identical residues in all the enzymes except for one. Dots indicate mismatches. The serine residue involved in the active site is indicated with *. The figure in the right hand corner shows how the domains are oriented.

Figure VII:

gives the amino acid sequence 1 - 722 for the entire protein (one letter code) and indicates exons a, b, c, and d. The sign # indicates the starting point of the repetition part.

Table 1. Amino acid sequence of BSSL peptides.

Due to interfering peaks no positive identification of the residue in cycles 1 and 2 of the sequencing could be made in peptide number 26. The peptide numbers refer to the peaks in Fig I.

Tryptic fragments	
TP16:	LysValThrGluGluAspPheTyrLys
TP19:	GlyIleProPheAlaAlaProThrLys
Tp20:	LeuValSerGluPheThrIleThrLys
TP24:	ThrTyrAlaTyrLeuPheSerHisProSerArg
TP26:	PheAspValTyrThrGluSerTrpAlaGlnAsp ProSerGlnGluAsnLys

Table 2: Identification of the exons a, b, c and d numbered as in Figures II and VII.

exon	Location between nucleotide number	between aminoacid number
a	986-1172	279-341
b	1173-1376	342-409
c	1377-1574	410-474
d	1575-2415	475-722
the entire protein	151-2316	1-722
the entire protein excluding repetitions	151-1755	1-535

References

1. Bläckberg, L., Ängquist, K.A, & Hernell, O. (1987) FEBS Lett. 217, 37-41.
2. Hernell, O. (1975) Eur. J. Clin. Invest. 5, 267-272.
3. Hernell, O., Bläckberg, L. & Bernbäck, S. (1988) In Perinatal nutrition (Lindblad, B.S., ed.) Bristol-Myers Nutrition symposia vol. 6, pp. 259272, Academic Press, New York.
4. Hernell, O, L. Bläckberg, B. Fredrikzon, and T. Olivecrona. 1981. Bile salt-stimulated lipase in human milk and lipid digestion in the neonatal period. In Textbook of gastroenterology and nutrition in infancy. E. Lebenthal, editor. Raven Press, New York. 465-471.
5. Bernbäck, S., Bläckberg, L. & Hernell, O. (1990) J. Clin. Invest. J. Clin. Invest. 221-226 (1990)
6. Bläckberg, L. & Hernell, O. (1981) Eur. J. Biochem 116, 221-225.
7. Hernell, O. & Bläckberg, L. (1982) Pediatr. Res. 16, 882-885.
8. Bläckberg, L. Lombardo, D., Hernell, O., Guy, O. & Olivecrona, T. (1981) FEBS Lett. 136, 284-288.
9. Fredrikzon, B., Hernell, O., Bläckberg, L. & Olivecrona, T. (1978) Pediatr. Res. 12, 1048-1052.
10. Wang, C.-S., Hartsuck, J.A. & Downs, D. (1988) Biochemistry 27, 4834-4840.
11. Bläckberg, L. & Hernell, O. (1983) FEBS Lett. 157, 337-341.
12. Björksten,B., Burman, L.G., deChateau, P., Fredrikzon, B., Gothefors, L. & Hernell, O. (1980) Br. Med. J. 201, 267-272.
13. Brueton, M.J., Berger, H.M., Brown, G.A., Ablitt, L., Iyangkaran, N. & Wharton B.A. (1978) Gut 19, 95-98.
14. Murphy, G.M. & Signer, E. (1975) Gut 15, 151-163.
15. Hernell, O., Staggers, J.E. & Carey, M.C. Biochemistry. (1990), 29:2041-2056.

16. Lombardo, D., Guy, O. & Figarella, C. (1978) Biochim. Biophys. Acta 527, 142-149.
17. Abouakil, N., Rogalska, E., Bonicel, J. & Lombardo, D. (1988) Biochim. Biophys. Acta 961, 299-308.
18. Hernell, O., Bläckberg, L. & Lindberg, T. (1988) in Textbook of Gastroenterology and Nutrition in infancy (Lebenthal, E., ed) pp. 209-217, Raven Press, New York.
19. Wang, C.-S. (1988) Biochem. Biophys. Res. Com. 155, 950-955.
20. Chirgwin, J.M. Przybyla, A.E., MacDonald R.J. & Rutter, W.J. (1979) Biochemistry 18, 5294-5299.
21. Aviv, H. & Leder, P. (1979) Proc. Natl. Acad. Sci. USA 69, 5201-5205.
22. Bailey, J.M. & Davidson, N. (1976) Anal. Biochem. 70, 75-85.
23. Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982): Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
24. Gubler, U. & Hoffman, B. J. (1983) Gene 25, 263-269.
25. Maniatis T., Haldison, R.C., Lacy, E., Lauer, J., O'Conell, C. & Qvon, D. (1978) Cell 15, 687-701.
26. Young, R.A. & Davis, R.W. (1983) Science 222, 778-782.
27. Young, R.A. & Davis, R.W. (1983) Proc. Natl. Acad. Sci, USA 80, 1194-1198.
28. McMaster, G.K. & Charmichael, G.G. (1977) Proc. Natl. Acad. Sci. USA 74, 48 353-48 358.
29. Thomas, P. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205.
30. Feinberg, A.P. & Vogelstein, B. (1983) Analyt. Biochem. 132, 6-13.
31. Sanger, F., Nicklen, S. & Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
32. Staden, R. (1984) Nucl, Acids Res. 12, 551-567.
33. Devereux, J., Haeberli, P. & Smithies, O. (1984) Nucl. Acids Res. 12, 387-395.

34. Wang, C.-S. & Johnson, K. (1983) Anal. Biochem. 133, 457-461.

35. Hamosh, M. (1986) in Human Milk Infant Nutrition and Health (Howell, R.R. Morriss, F.H. & Pickering, L.K. eds) pp. 66-97, Charles, C. Thomas, Springfield.

36. Wang, C.-S. (1980) Anal. Biochem. 105, 398-402.

37. Wang, C.-S & Lee, D.M. (1985) J. Lipid. Res. 26, 824-830.

38. Brenner, S. (1988) Nature 334, 528-530.

39. Yang, C.-Y., Gu, Z.-W., Yang, H.-H., Rohde, M.F., Gotto, Jr., A.M. & Pownall, H.J. (1989) J. Biol. Chem. 265, 16822-16827.

40. Han, J.H., Stratowa, C., & Rutter, W.J. (1987) Biochemistry 26, 1617-1625.

41. Kyger, E., Wiegand, R., & Lange, L. (1989) Biochim. Biophys. Res. Com 164, 1302-1309.

42. Bläckberg, L. (1981) Fat digestion in newborn infant. Umeå University Medical Dissertations New Series No 71.

43. Rudd, E.A. & Brockman, H.L. (1984) in Lipases (Borgström, B & Brockman, H.L. eds) pp. 184-204, Elsevier, Amsterdam.

44. Mickel, S., Weidenbach, F., Swarovsky, B., LaForge, S. & Scheele, G. (1989) J. Biol. Chem. 264, 12895-12901.

45. Cammulli, E.D., Linke, M.J., Brockman, H.L. & Hui, D.Y. (1989) Biochim. Biophys. Acta 1005, 177-182.

46. Mann, S.K.O. & Firtel, R.A. (1987) Mol. Cell Biol. 7, 458-469.

47. Mercken, L., Simons, M.J., Swillens, S., Masser, M. & Vassart, G. (1985) Nature 316, 647-651.

48. Lauro, R., Obici, S. Condliffe, D., Ursini, V.M., Musti, A., Moscatelli, C. & Avvedimento, V.E. (1985) Eur. J. Biochem. 148, 7-11.

49. Malthiery, Y. & Lissitzky, S. (1987) Eur, J. Biochem. 165, 491-498.

50. Prody, C., Zevin-Sonkin, D., Gnatt, A., Goldberg, O.,

& Soreq, H. (1987) Proc. Natl. Acad. Sci. USA 84, 3555-3559.

51. Sikorav, J-L., Krejci, E. & Massoulie', J. (1987) EMBO J. 6, 1865-1873.

52. Oakeshott, J.G. Collet, C., Phillis, R.W., Nielsen, K.M., Russel, R.J., Cambers, G.K., Ross, V. & Richmond, R.C. (1987) Proc. Natl. Acad, Sci, USA 84, 3359-3363.

53. Long, R., Satho, H., Martin, B., Kimura, S., Gonzalez, F. & Pohl, L. (1988) Biochem. Biophys. Res. Comm. 156, 866-873.

54. Hall, L.M.C. & Spierer, P. (1986) EMBO J. 5, 2949-2954.

55. Sjöberg, S., Carlsson, P., Enerbäck, S., & Bjursell, G. (1989) CABIOS 5, 41-46.

56. EP-A-317355.

57. Hernell, O., and T. Olivecrona. 1974. Human milk lipases II. Bile salt-stimulated lipase. Biochim. Biophys. Acta 369:234-244.

58. Hernell, O., L. Bläckberg, and T. Olivecrona. 1981. Human milk lipases. In Textbook of gastroenterology and nutrition in infancy. E. Lebenthal, editor. Raven Press, New York. 347-354.

59. Atkinson, S.A., M.H. Bryan, and G.H. Andersson. 1981. Human milk feeding in premature infants: protein, fat and carbohydrate balances in the first two weeks of life. J.Pediatr. 99:617-624.

60. Chappell, J.E., M.T. Clandinin, C. Kearney-Volpe, B. Reichman and P.W. Swyer. 1986. Fatty acid balance studies in premature infants fed human milk or formula: effect of calcium supplementation. J. Pediatr. 108:438-447.

61. Williamson, S., E. Finucane, H. Ellis, and H.R. Gamsu. 1978. Effect of heat treatment of human milk on absorption of nitrogen, fat, sodium, calcium and phosphorus by preterm infants. Arch. Dis. Childhood 53:555-563.

What we claim is:

1. A protein as indicated in Figure VII from position 1 to position 722.

2. A protein as indicated in Figure VII from position 1 to position 535.

3. A protein as indicated in Figure VII from position 1 to position 278.

4. A protein as indicated in Figure VII from position 1 to position 341.

5. A protein as indicated in Figure VII from position 1 to position 409.

6. A protein as indicated in Figure VII from position 1 to position 474.

7. A protein as indicated in Figure VII from position 1 to position 278, position 279 to position 341, position 279 to position 409, position 279 to position 474, position 342 to position 409, position 342 to position 474, or position 536 to position 722.

8. A protein according to any of claims 1-7 in combination with one or more of the repeats according to Figure V.

9. A protein according to any of claims 1-8, possessing methionine as additional N-terminal amino acid.

10. A protein according to any of claims 1-9, exhibiting one or several of the critical functions of naturally occurring bile salt-stimulated lipase.

11. A functionally equivalent variant or mutant of a protein according to any of claims 1-9.

12. A DNA sequence coding for a protein with the amino acid sequence according to claim 1.

13. A DNA sequence coding for a protein with the amino acid sequence according to claim 2.

14. A DNA sequence coding for a protein with the amino acid sequence according to any of claims 3, 4, 5, 6, 7, 8, 9, 10 and 11.

15. A DNA sequence defined by the following nucleotide numbers in Figure II:

151-2316
151-1755
151-985
151-1172
151-1376
151-1574
986-1172
986-1376
986-1574
1173-1376
1173-1574

16. A vector comprising a DNA sequence coding for a protein according to claims 1 - 11

17. A vector comprising a DNA sequence according to claims 12 - 15.

18. A host organism transformed with a vector according to claims 16 or 17.

19. A process for the production of a protein according to claims 1 - 11 by growing a host organism containing a vector according to claims 16 or 17, and isolating the protein.

20. A pharmaceutical composition comprising a protein according to any of claims 1 - 11.

21. A pharmaceutical composition comprising a protein according to any of claims 1-11 in combination with a lipase or in combination with preparations containing lipases.

22. The use of a protein according to any of claims 1 - 11, for the manufacture of a medicament for the treatment of a pathological condition related to exocrine pancreatic insufficiency.

23. The use of a protein according to any of claims 1 - 11 for the manufacture of a medicament for the treatment of cystic fibrosis.

24. The use of a protein according to any of claims 1 - 11 as a supplement to an infant food formulation.

25. The use of a protein according to any of claims 1-11 for the manufacture of a medicament for the treatment of chronic pancreatitis.

26. The use of a protein according to any of claims 1-11 for the manufacture of a medicament for the treatment of fat malabsorption.

27. The use of a protein accordin to any of claims 1-11 for the manufacture of a medicament for the treatment of malabsorption of fat soluble vitamins.

28. The use of a protein according to any of claims 1-11 for the manufacture of a medicament for the treatment of fat malabsorption due to physiological reasons.

29. The use, according to claims 22-28 of a protein according to any of claims 1-11 in combination with a lipase or lipases or in combination with preparations containing a lipase or lipases.

30. An infant food formulation, supplemented with a protein according to any of claims 1 - 11.

31. A protein according to claims 1 - 11 in substantially pure form.

32. A protein according to claims 1 - 11 in isolated form.

33. A DNA sequence according to claims 12 - 15 in substantially pure form.

34. A DNA sequence according to claims 12 - 15 in isolated form.

35. A process for the preparation of a pharmaceutical composition according to claims 20 or 21 by incorporating a protein according to claims 1 - 11, 31, or 32 in a pharmacologically acceptable carrier.

36. A process according to claim 35 for the preparation of a pharmaceutical composition for oral administration.

37. A process for the preparation of an infant food formulation according to claim 30, by supplementing an infant food formulation with a protein according to claims 1 -11, 31, or 32.

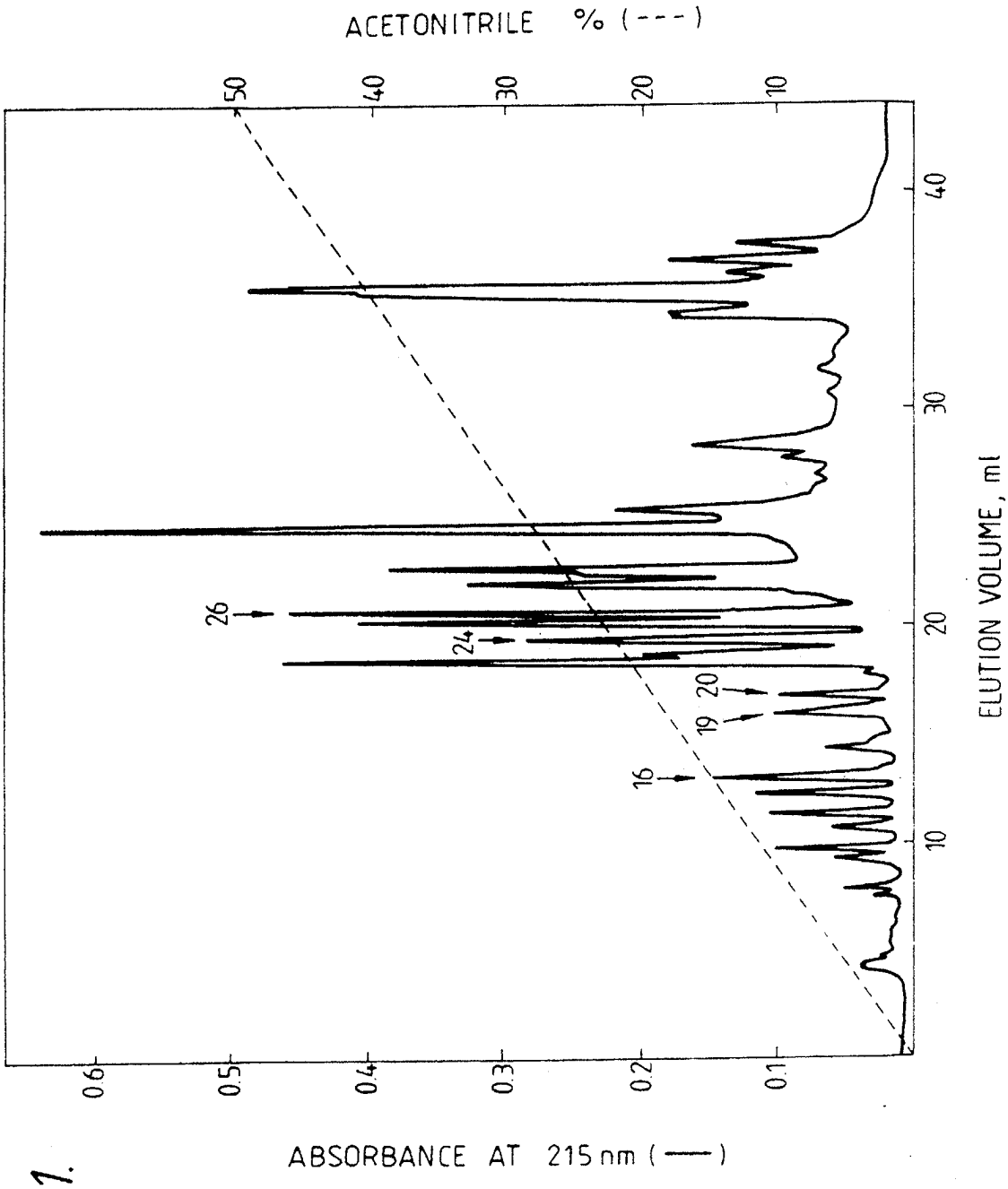
ACETONITRILE % (---)
50
40
30
20
10
40
30
20
10
ELUTION VOLUME, ml
26
24
20
19
16
0.6
0.5
0.4
0.3
0.2
0.1
ABSORBANCE AT 215 nm (—)

Fig. 1.

Fig. 2 (1/4)

```
ACCTTCTGTA TCAGTTAAGT GTCAAGATGG AAGGAACAGC AGTCTCAAGA TAATGCAAAG      60

AGTTTATTCA TCCAGAGGCT G ATG CTC ACC ATG GGG CGC CTG CAA CTG GTT      111
                        Met Leu Thr Met Gly Arg Leu Gln Leu Val
                          1               5                  10

GTG TTG GGC CTC ACC TGC TGC TGG GCA GTG GCG AGT GCC GCG AAG CTG      159
Val Leu Gly Leu Thr Cys Cys Trp Ala Val Ala Ser Ala Ala Lys Leu
                  15                  20                  25

GGC GCC GTG TAC ACA GAA GGT GGG TTC GTG GAA GGC GTC AAT AAG AAG      207
Gly Ala Val Tyr Thr Glu Gly Gly Phe Val Glu Gly Val Asn Lys Lys
              30                  35                  40

CTC GGC CTC CTG GGT GAC TCT GTG GAC ATC TTC AAG GGC ATC CCC TTC      255
Leu Gly Leu Leu Gly Asp Ser Val Asp Ile Phe Lys Gly Ile Pro Phe
          45                  50                  55

GCA GCT CCC ACC AAG GCC CTG GAA AAT CCT CAG CCA CAT CCT GGC TGG      303
Ala Ala Pro Thr Lys Ala Leu Glu Asn Pro Gln Pro His Pro Gly Trp
      60                  65                  70

CAA GGG ACC CTG AAG GCC AAG AAC TTC AAG AAG AGA TGC CTG CAG GCC      351
Gln Gly Thr Leu Lys Ala Lys Asn Phe Lys Lys Arg Cys Leu Gln Ala
  75                  80                  85                  90

ACC ATC ACC CAG GAC AGC ACC TAC GGG GAT GAA GAC TGC CTG TAC CTC      399
Thr Ile Thr Gln Asp Ser Thr Tyr Gly Asp Glu Asp Cys Leu Tyr Leu
                  95                 100                 105

AAC ATT TGG GTG CCC CAG GGC AGG AAG CAA GTC TCC CGG GAC CTG CCC      447
Asn Ile Trp Val Pro Gln Gly Arg Lys Gln Val Ser Arg Asp Leu Pro
              110                 115                 120

GTT ATG ATC TGG ATC TAT GGA GGC GCC TTC CTC ATG GGG TCC GGC CAT      495
Val Met Ile Trp Ile Tyr Gly Gly Ala Phe Leu Met Gly Ser Gly His
          125                 130                 135

GGG GCC AAC TTC CTC AAC AAC TAC CTG TAT GAC GGC GAG GAG ATC GCC      543
Gly Ala Asn Phe Leu Asn Asn Tyr Leu Tyr Asp Gly Glu Glu Ile Ala
      140                 145                 150

ACA CGC GGA AAC GTC ATC GTG GTC ACC TTC AAC TAC CGT GTC GGC CCC      591
Thr Arg Gly Asn Val Ile Val Val Thr Phe Asn Tyr Arg Val Gly Pro
155                 160                 165                 170

CTT GGG TTC CTC AGC ACT GGG GAC GCC AAT CTG CCA GGT AAC TAT GGC      639
Leu Gly Phe Leu Ser Thr Gly Asp Ala Asn Leu Pro Gly Asn Tyr Gly
                  175                 180                 185
```

Fig. 2 (2/4)

```
CTT CGG GAT CAG CAC ATG GCC ATT GCT TGG GTG AAG AGG AAT ATC GCG        687
Leu Arg Asp Gln His Met Ala Ile Ala Trp Val Lys Arg Asn Ile Ala
            190                 195                 200

GCC TTC GGG GGG GAC CCC AAC AAC ATC ACG CTC TTC GGG GAG TCT GCT        735
Ala Phe Gly Gly Asp Pro Asn Asn Ile Thr Leu Phe Gly Glu Ser Ala
        205                 210                 215

GGA GGT GCC AGC GTC TCT CTG CAG ACC CTC TCC CCC TAC AAC AAG GGC        783
Gly Gly Ala Ser Val Ser Leu Gln Thr Leu Ser Pro Tyr Asn Lys Gly
    220                 225                 230

CTC ATC CGG CGA GCC ATC AGC CAG AGC GGC GTG GCC CTG AGT CCC TGG        831
Leu Ile Arg Arg Ala Ile Ser Gln Ser Gly Val Ala Leu Ser Pro Trp
235                 240                 245                 250

GTC ATC CAG AAA AAC CCA CTC TTC TGG GCC AAA AAG GTG GCT GAG AAG        879
Val Ile Gln Lys Asn Pro Leu Phe Trp Ala Lys Lys Val Ala Glu Lys
                255                 260                 265

GTG GGT TGC CCT GTG GGT GAT GCC GCC AGG ATG GCC CAG TGT CTG AAG        927
Val Gly Cys Pro Val Gly Asp Ala Ala Arg Met Ala Gln Cys Leu Lys
            270                 275                 280

GTT ACT CAT CCC CGA GCC CTG ACG CTG GCC TAT AAG GTG CCG CTG GCA        975
Val Thr His Pro Arg Ala Leu Thr Leu Ala Tyr Lys Val Pro Leu Ala
        285                 290                 295
             *
GGC CTG GAG TAC CCC ATG CTG CAC TAT GTG GGC TTC GTC CCT GTC ATT       1023
Gly Leu Glu Tyr Pro Met Leu His Tyr Val Gly Phe Val Pro Val Ile
    300                 305                 310

GAT GGA GAC TTC ATC CCC GCT GAC CCG ATC AAC CTG TAC GCC AAC GCC       1071
Asp Gly Asp Phe Ile Pro Ala Asp Pro Ile Asn Leu Tyr Ala Asn Ala
315                 320                 325                 330

GCC GAC ATC GAC TAT ATA GCA GGC ACC AAC AAC ATG GAC GGC CAC ATC       1119
Ala Asp Ile Asp Tyr Ile Ala Gly Thr Asn Asn Met Asp Gly His Ile
                335                 340                 345

TTC GCC AGC ATC GAC ATG CCT GCC ATC AAC AAG GGC AAC AAG AAA GTC       1167
Phe Ala Ser Ile Asp Met Pro Ala Ile Asn Lys Gly Asn Lys Lys Val
            350                 355                 360
      *
ACG GAG GAG GAC TTC TAC AAG CTG GTC AGT GAG TTC ACA ATC ACC AAG       1215
Thr Glu Glu Asp Phe Tyr Lys Leu Val Ser Glu Phe Thr Ile Thr Lys
        365                 370                 375

GGG CTC AGA GGC GCC AAG ACG ACC TTT GAT GTC TAC ACC GAG TCC TGG       1263
Gly Leu Arg Gly Ala Lys Thr Thr Phe Asp Val Tyr Thr Glu Ser Trp
    380                 385                 390
```

Fig. 2 (3/4)

```
GCC CAG GAC CCA TCC CAG GAG AAT AAG AAG AAG ACT GTG GTG GAC TTT      1311
Ala Gln Asp Pro Ser Gln Glu Asn Lys Lys Lys Thr Val Val Asp Phe
395                 400                 405                 410

GAG ACC GAT GTC CTC TTC CTG GTG CCC ACC GAG ATT GCC CTA GCC CAG      1359
Glu Thr Asp Val Leu Phe Leu Val Pro Thr Glu Ile Ala Leu Ala Gln
                415                 420                 425
                    *
CAC AGA GCC AAT GCC AAG AGT GCC AAG ACC TAC GCC TAC CTG TTT TCC      1407
His Arg Ala Asn Ala Lys Ser Ala Lys Thr Tyr Ala Tyr Leu Phe Ser
            430                 435                 440

CAT CCC TCT CGG ATG CCC GTC TAC CCC AAA TGG GTG GGG GCC GAC CAT      1455
His Pro Ser Arg Met Pro Val Tyr Pro Lys Trp Val Gly Ala Asp His
        445                 450                 455

GCA GAT GAC ATT CAG TAC GTT TTC GGG AAG CCC TTC GCC ACC CCC ACG      1503
Ala Asp Asp Ile Gln Tyr Val Phe Gly Lys Pro Phe Ala Thr Pro Thr
    460                 465                 470

GGC TAC CGG CCC CAA GAC AGG ACA GTC TCT AAG GCC ATG ATC GCC TAC      1551
Gly Tyr Arg Pro Gln Asp Arg Thr Val Ser Lys Ala Met Ile Ala Tyr
475                 480                 485                 490
                            *
TGG ACC AAC TTT GCC AAA ACA GGG GAC CCC AAC ATG GGC GAC TCG GCT      1599
Trp Thr Asn Phe Ala Lys Thr Gly Asp Pro Asn Met Gly Asp Ser Ala
                495                 500                 505

GTG CCC ACA CAC TGG GAA CCC TAC ACT ACG GAA AAC AGC GGC TAC CTG      1647
Val Pro Thr His Trp Glu Pro Tyr Thr Thr Glu Asn Ser Gly Tyr Leu
            510                 515                 520

GAG ATC ACC AAG AAG ATG GGC AGC AGC TCC ATG AAG CGG AGC CTG AGA      1695
Glu Ile Thr Lys Lys Met Gly Ser Ser Ser Met Lys Arg Ser Leu Arg
        525                 530                 535

ACC AAC TTC CTG CGC TAC TGG ACC CTC ACC TAT CTG GCG CTG CCC ACA      1743
Thr Asn Phe Leu Arg Tyr Trp Thr Leu Thr Tyr Leu Ala Leu Pro Thr
    540                 545                 550
                #
GTG ACC GAC CAG GAG GCC ACC CCT GTG CCC CCC ACA GGG GAC TCC GAG      1791
Val Thr Asp Gln Glu Ala Thr Pro Val Pro Pro Thr Gly Asp Ser Glu
555                 560                 565                 570

GCC ACT CCC GTG CCC CCC ACG GGT GAC TCC GAG ACC GCC CCC GTG CCG      1839
Ala Thr Pro Val Pro Pro Thr Gly Asp Ser Glu Thr Ala Pro Val Pro
                575                 580                 585

CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC      1887
Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser
            590                 595                 600
```

Fig. 2 (4/4)

```
GGG GCC CCC CCC TTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG      1935
Gly Ala Pro Pro Leu Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val
        605                 610                 615

CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC      1983
Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp
    620                 625                 630

TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC      2031
Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro
635                 640                 645                 650

GTG CCG CCC ACG GGT GAC TCC GGC GCC CCC CCC GTG CCG CCC ACG GGT      2079
Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly
                655                 660                 665

GAC GCC GGG CCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGC GCC CCC      2127
Asp Ala Gly Pro Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro
            670                 675                 680

CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG ACC CCC ACG      2175
Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Thr Pro Thr
        685                 690                 695

GGT GAC TCC GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC      2223
Gly Asp Ser Glu Thr Ala Pro Val Pro Pro Thr Gly Asp Ser Gly Ala
    700                 705                 710

CCC CCT GTG CCC CCC ACG GGT GAC TCT GAG GCT GCC CCT GTG CCC CCC      2271
Pro Pro Val Pro Pro Thr Gly Asp Ser Glu Ala Ala Pro Val Pro Pro
715                 720                 725                 730

ACA GAT GAC TCC AAG GAA GCT CAG ATG CCT GCA GTC ATT AGG TTT          2316
Thr Asp Asp Ser Lys Glu Ala Gln Met Pro Ala Val Ile Arg Phe
                735                 740                 745

TAGCGTCCCA TGAGCCTTGG TATCAAGAGG CCACAAGAGT GGGACCCCAG GGGCTCCCCT    2376

CCCATCTTGA GCTCTTCCTG AATAAAGCCT CATACCCCTA AAAAAAAAAA AA            2428
```

Fig. 3.

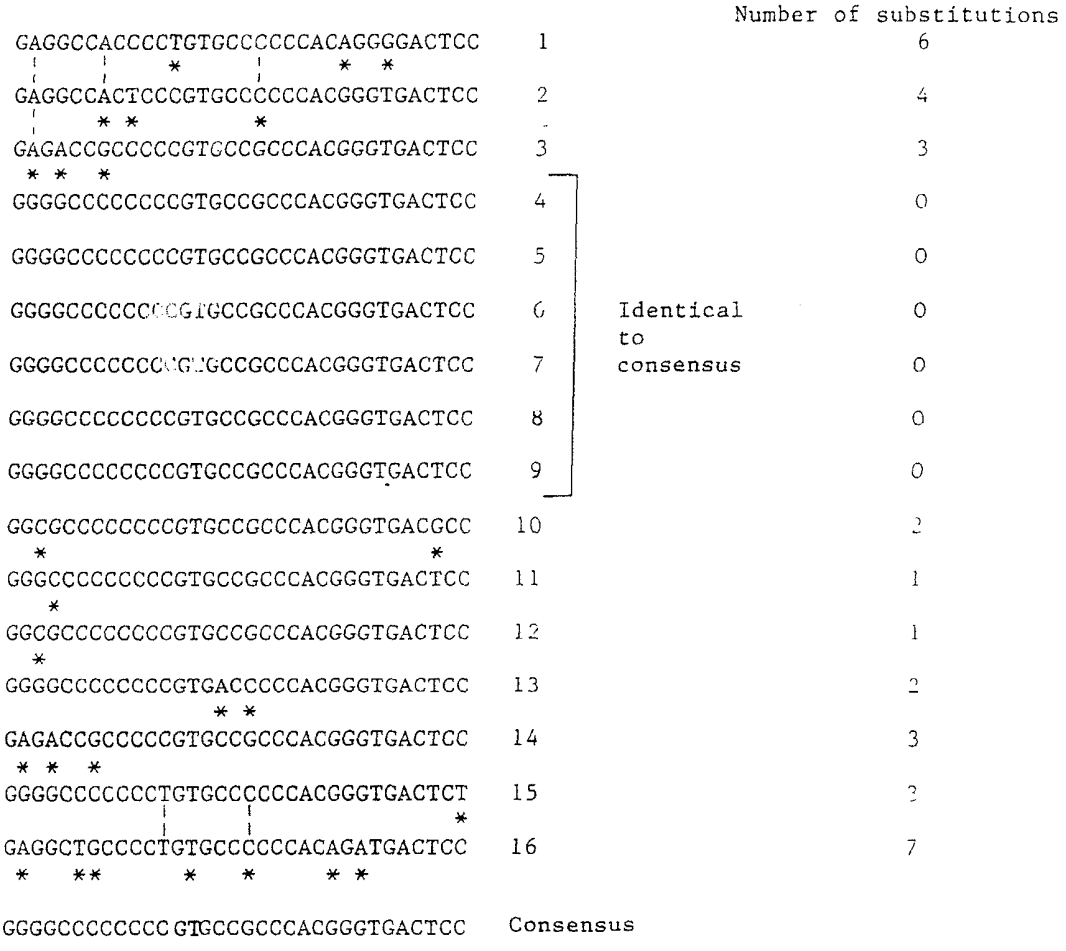
Number of substitutions
GAGGCCACCCCTGTGCCCCCCACAGGGGACTCC 1 6
GAGGCCACTCCCGTGCCCCCCACGGGTGACTCC 2 4
GAGACCGCCCCCCGTGCCGCCCACGGGTGACTCC 3 3
GGGGCCCCCCCCCGTGCCGCCCACGGGTGACTCC 4 0
GGGGCCCCCCCCCGTGCCGCCCACGGGTGACTCC 5 0
GGGGCCCCCCCCGTGCCGCCCACGGGTGACTCC 6 0
GGGGCCCCCCCCGTGCCGCCCACGGGTGACTCC 7 0
GGGGCCCCCCCCCGTGCCGCCCACGGGTGACTCC 8 0
GGGGCCCCCCCCCGTGCCGCCCACGGGTGACTCC 9 0
Identical to consensus
GGCGCCCCCCCCCGTGCCGCCCACGGGTGACGCC 10 2
GGGCCCCCCCCCCGTGCCGCCCACGGGTGACTCC 11 1
GGCGCCCCCCCCCGTGCCGCCCACGGGTGACTCC 12 1
GGGGCCCCCCCCCCGTGACCCCCACGGGTGACTCC 13 2
GAGACCGCCCCCGTGCCGCCCACGGGTGACTCC 14 3
GGGGCCCCCCCCTGTGCCCCCCACGGGTGACTCT 15 3
GAGGCTGCCCCTGTGCCCCCCACAGATGACTCC 16 7
GGGGCCCCCCCCGTGCCGCCCACGGGTGACTCC Consensus

Fig. 4.

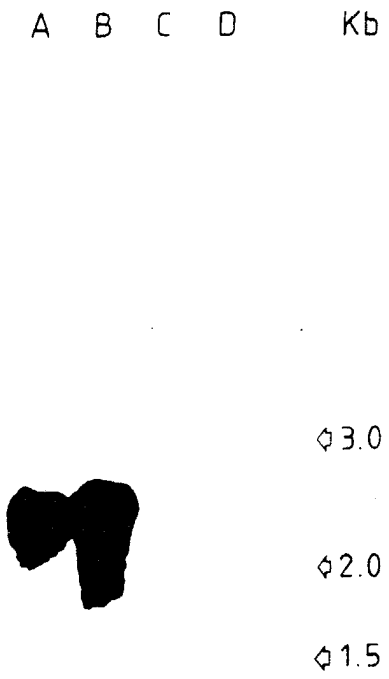

Fig. 5 (1/3)

```
                                                                    50
Bssl        MLTMGRLQLVVLGLTCCWAVASAAKLGAVYTEGGFVEGVNKKLGLLG.DS
Ratlpl      ...MGRLEVLFLGLTCCLAAACAAKLGALYTEGGFVEGVNKKLSLLGGDS
Bovceh      ....MLGASRLGPSPGCLAVASAAKLGSVYTEGGFVEGVNKKLSLFG.DS

Consensus   ....grl....lgltcclaaaaAAKLGavYTEGGFVEGVNKKLsLlG.DS


                                                                   100
Bssl        VDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNFKKRCLQATITQDSTYG
Ratlpl      VDIFKGIPFA.TAKTLENPQRHPGWQGTLKATQFKKRCLQATITQDDTYG
Bovech      VDIFKGIPFAAAPKALEKPKRHPGWQGTLKAKSFKKRCLQATLTQDSTYG

Consensus   VDIFKGIPFAa..KaLEnPqrHPGWQGTLKAk.FKKRCLQATiTQDsTYG


                                                                   150
Bssl        DEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGBGANFLNNYLYD
Ratlpl      QEDCLYLNIWVPQGRKQVSHDLPVMVWIYGGAFLMGSGQGANFLKNYLYD
Bovceh      NEDCLYLNIWVPQGRKEVSHDLPVMIWIYGGAFLMGASQGANFLSNYLYD

Consensus   .EDCLYLNIWVPQGRKqVShDLPVMiWIYGGAFLMGsgqGANFL.NYLYD


                                                                   200
Bssl        GKKIATRGNVIVVTFNYRVGPLGFLSTGDANLPGNYGLRDQHMAIAWVKR
Ratlpl      GKKIATRANVIVVTFNYRVGPLGFLSTGDANLPGNFGLRDQHMAIAWVKR
Bovech      GKKIATRGNVIVVTFNYRVGPLGFLSTGDSNLPGNYGLWDQHMAIAWVKR

Consensus   GKKIATRgNVIVVTFNYRVGPLGFLSTGDaNLPGNyGLrDQHMAIAWVKR


                      #       *                                    250
Bssl        NIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSP
Ratlpl      NIAAFGGDPNNITIFGESAGGAIVSLQTLSPYNKGLIRRAISQSGVALSP
Bovceh      NIEAFGGDPDNITLFGESAGGASVSLQTLSPYNKGLIKRGISQSGVGLCP

Consensus   NIaAFGCDPDNITLFGESAGGAsVSLQTLSPYNKGLIrRaISQSGVaLsP


                                                                   300
Bssl        WVIQKNPLFWAKKVAKKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAG
Ratlpl      WAIQENPLFWAKTIAKKVGCPTEDTAKMAGCLKITDPRALTLAYRLPLKS
Bovech      WAIQQDPLFWAKRIAKKVGCPVDDTSKMAGCAKITDPRALTLAYKLPLGS

Consensus   WaIQ.nPLFWAK.iAKKVGCPv.DtakMAgClKiTDPRALTLAYklPL.s


                                                                   350
Bssl        LEYPMLHYVGFVPVIDGDFIPADPINLYANAADIDYIAGTNNMDGHIFAS
Ratlpl      QEYPIVEYLAFIPVVDGDFIPDDPINLYDNAADIDYLAGINDMDGHLFAT
Bovceh      TEYPKLHYLSFVPVIDGDFIPDDPVNLYANAADVDYIAGTNDMDGHLFVG

Consensus   .EYP.lHYl.FvPViDGDFIPdDPiNLYaNAADiDYiAGtNdMDGHlFa.
```

Fig. 5 (2/3)

```
                                                                    400
Bssl        IDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGAKTTFDVYTESWAQDPS
Ratlpl      VDVPAIDKAKQDVTEEDFYRLVSGHTVAKGLKGTQATFDIYTESWAQDPS
Bovech      MDVPAINSNKQDVTEEDFYKLVSGLTVTKGLRGANATYEVYTEPWAQDSS

Consensus   .DvPAInk.kqdVTEEDFYkLVSg.TvtKGLrGa.aTfdvYTEsWAQDpS

                                                                    450
Bssl        QENKKKTVVDFETDVLFLVPTEIALAQHRANAKSAKTYAYLFSHPSRMPV
Ratlpl      QENMKKTVVAFETDILFLIPTEMALAQHRAHAKSAKTYSYLFSHPSRMPI
Bovceh      QETRKKTMVDLETDILFLIPTKIAVAQHKSHAKSANTYTYLFSQPSRMPI

Consensus   QEn.KKTvVdfETDiLFLiPTeiAlAQHrahAKSAkTY.YLFShPSRMPi

                                                                    500
Bssl        YPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGD
Ratlpl      YPKWMGADHADDLQYVFGKPFATPLGYRAQDRTVSKAMIAYWTNFAKSGD
Bovech      YPKWMGADHADDLQYVFGKPFATPLGYRAQDRTVSKAMIAYWTNFARTGD

Consensus   YPKWmGADHADDlQYVFGKPFATPlGYRaQDRTVSKAMIAYWTNFAktGD

                                                                    550
Bssl        PNMGDSAVPTHWEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLTY
Ratlpl      PNMGNSPVPTHWYPYTMENGNYLDINKKITSTSMKEHLREKFLKFWAVTF
Bovceh      PNTGHSTVPANWDPYTLEDDNYLEINKQMDSNSMKLRTLTNYLQFWTQTY

Consensus   PNnG.S.VPth..PYT.Kn.nYleInKkm.S.SMK.hLRtnfL.fWt.Ty

                                                                    596
Bssl        LALPTVTDQ EATPVPPTGDS EATPVPPTGDS ETAPVPPTGDS GAPP
Ratlpl      EMLPTV...  VGDHTPPKDDS EAAPVPPTDDS DGGPVPPTDDS QTTP
Bovech      QALPTVTSA  GASLLPPEDNS QASPVPPADNS GAPTEPSAGDS ....

Consensus   .aLPTVt....a...PP.ddS.eA.PVPPtddS....pvPptgDS....p

                                                                    642
Bssl        VPPTGDS GAPPVPPTGDS GAPPVPPTGDS GAPPVPPTGDS GAPPVP
Ratlpl      VPPTDNS QA..........  ..........  ..........  ......
Bovceh      .......  ..........  ..........  ..........  ......

Consensus   ......s..a.....................................
```

Fig. 5 (3/3)

```
                                                                                                  689
Bssl        PTGDS | GAPPVPPTGDS | GAPPVPPTGDS | GPPPVTPTGDS | GAPPVPPTG
Rarlpl      ..... | ........... | ........... | ........... | .........
Bovech      ..... | ........... | ........... | ........... | .........

Consensus   .....................................................

                                                                                                  735
Bssl        DS | GAPPVTPTGDS | ETAPVPPTGDS | GAPPVPPTGD | SEAAPVPPTDDS
Ratlpl      .. | ........... | ........... | ......... | .........GDS
Bovceh      .. | ........... | ........... | ......... | ...........

Consensus   .....................................................ds


Bssl        KE.AQMPAVIRF
Ratlpl      VE.AQMPGPIGF
Bovech      .EVAQMPVVIGF

Consensus   .e.AQMP.vIgF
```

Fig. 6 (1/3)

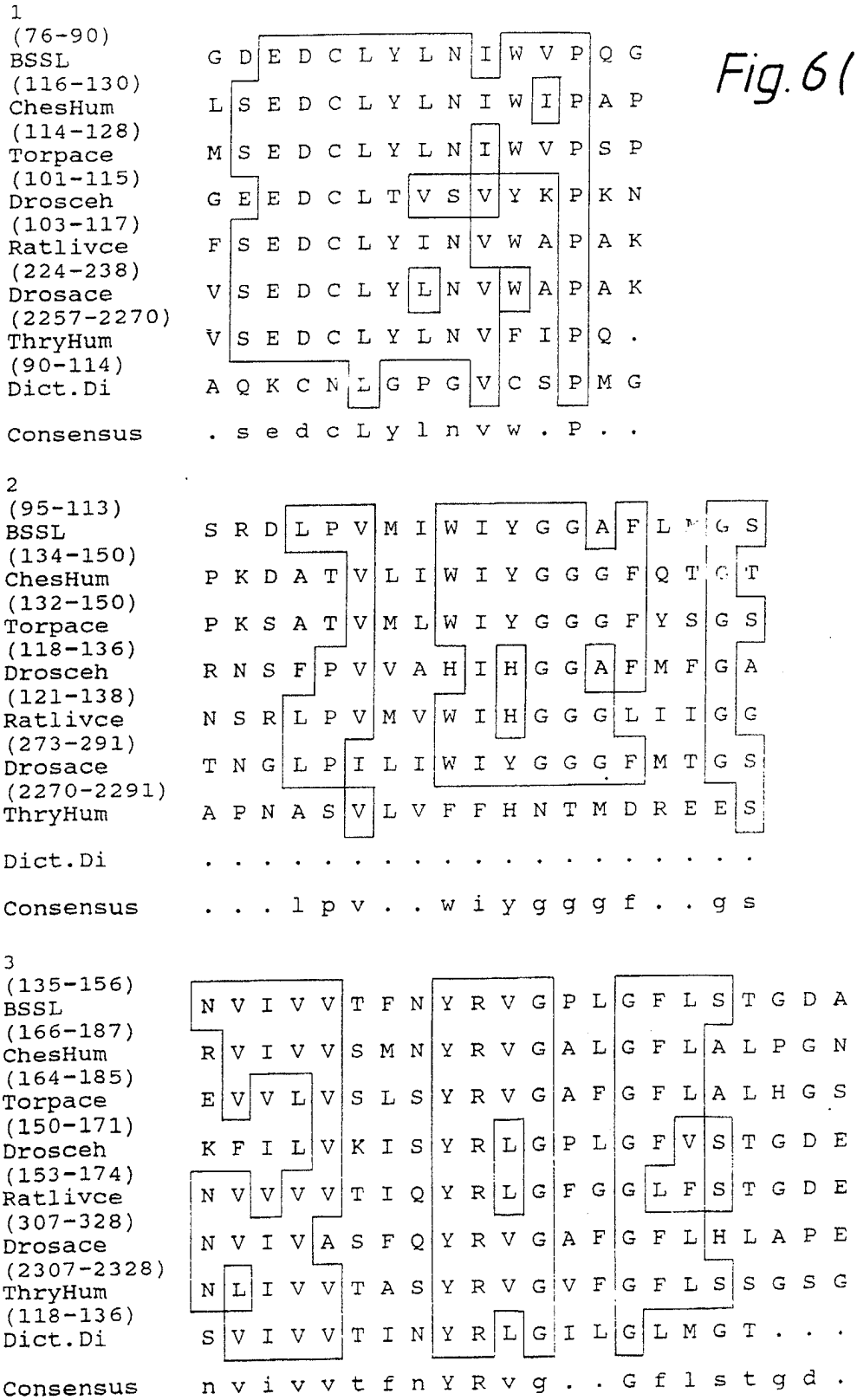
1
(76-90)
BSSL G D E D C L Y L N I W V P Q G
(116-130)
ChesHum L S E D C L Y L N I W I P A P
(114-128)
Torpace M S E D C L Y L N I W V P S P
(101-115)
Drosceh G E E D C L T V S V Y K P K N
(103-117)
Ratlivce F S E D C L Y I N V W A P A K
(224-238)
Drosace V S E D C L Y L N V W A P A K
(2257-2270)
ThryHum V S E D C L Y L N V F I P Q .
(90-114)
Dict.Di A Q K C N L G P G V C S P M G
Consensus . s e d c L y l n v w . P . .
2
(95-113)
BSSL S R D L P V M I W I Y G G A F L M G S
(134-150)
ChesHum P K D A T V L I W I Y G G G F Q T G T
(132-150)
Torpace P K S A T V M L W I Y G G G F Y S G S
(118-136)
Drosceh R N S F P V V A H I H G G A F M F G A
(121-138)
Ratlivce N S R L P V M V W I H G G G L I I G G
(273-291)
Drosace T N G L P I L I W I Y G G G F M T G S
(2270-2291)
ThryHum A P N A S V L V F F H N T M D R E E S
Dict.Di . . . . . .
Consensus . . . l p v . . w i y g g g f . . g s
3
(135-156)
BSSL N V I V V T F N Y R V G P L G F L S T G D A
(166-187)
ChesHum R V I V V S M N Y R V G A L G F L A L P G N
(164-185)
Torpace E V V L V S L S Y R V G A F G F L A L H G S
(150-171)
Drosceh K F I L V K I S Y R L G P L G F V S T G D E
(153-174)
Ratlivce N V V V V T I Q Y R L G F G G L F S T G D E
(307-328)
Drosace N V I V A S F Q Y R V G A F G F L H L A P E
(2307-2328)
ThryHum N L I V V T A S Y R V G V F G F L S S G S G
(118-136)
Dict.Di S V I V V T I N Y R L G I L G L M G T . . .
Consensus n v i v v t f n Y R v g . . G f l s t g d .

Fig. 6 (2/3)

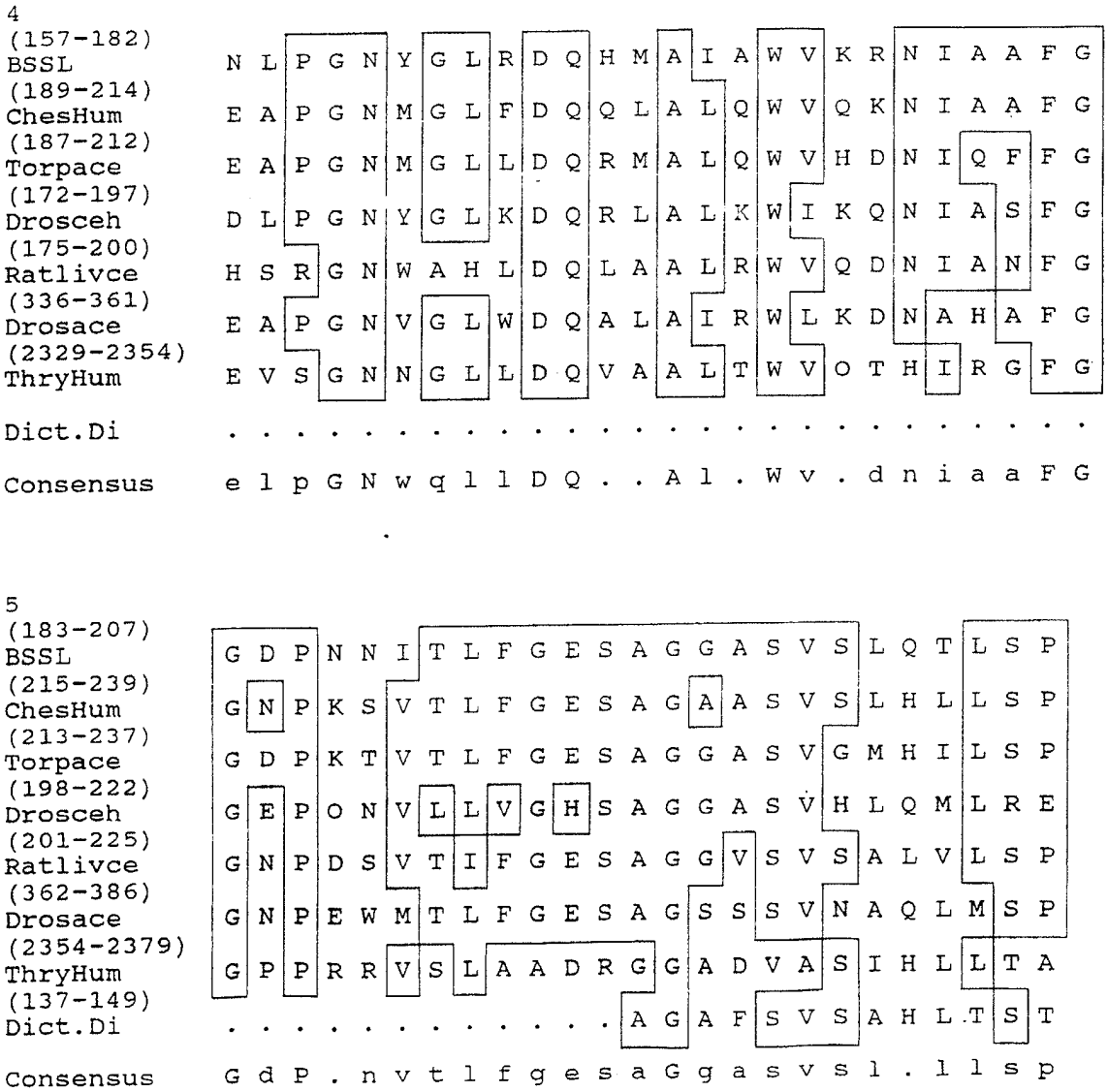
4
(157-182)
BSSL N L P G N Y G L R D Q H M A I A W V K R N I A A F G
(189-214)
ChesHum E A P G N M G L F D Q Q L A L Q W V Q K N I A A F G
(187-212)
Torpace E A P G N M G L L D Q R M A L Q W V H D N I Q F F G
(172-197)
Drosceh D L P G N Y G L K D Q R L A L K W I K Q N I A S F G
(175-200)
Ratlivce H S R G N W A H L D Q L A A L R W V Q D N I A N F G
(336-361)
Drosace E A P G N V G L W D Q A L A I R W L K D N A H A F G
(2329-2354)
ThryHum E V S G N N G L L D Q V A A L T W V O T H I R G F G
Dict.Di . . . . . .
Consensus e l p G N w q l l D Q . . A l . W v . d n i a a F G
5
(183-207)
BSSL G D P N N I T L F G E S A G G A S V S L Q T L S P
(215-239)
ChesHum G N P K S V T L F G E S A G A A S V S L H L L S P
(213-237)
Torpace G D P K T V T L F G E S A G G A S V G M H I L S P
(198-222)
Drosceh G E P O N V L L V G H S A G G A S V H L Q M L R E
(201-225)
Ratlivce G N P D S V T I F G E S A G G V S V S A L V L S P
(362-386)
Drosace G N P E W M T L F G E S A G S S S V N A Q L M S P
(2354-2379)
ThryHum G P P R R V S L A A D R G G A D V A S I H L L T A
(137-149)
Dict.Di . . . . . A G A F S V S A H L T S T
Consensus G d P . n v t l f g e s a G g a s v s l . l l s p

Fig.6 (3/3)

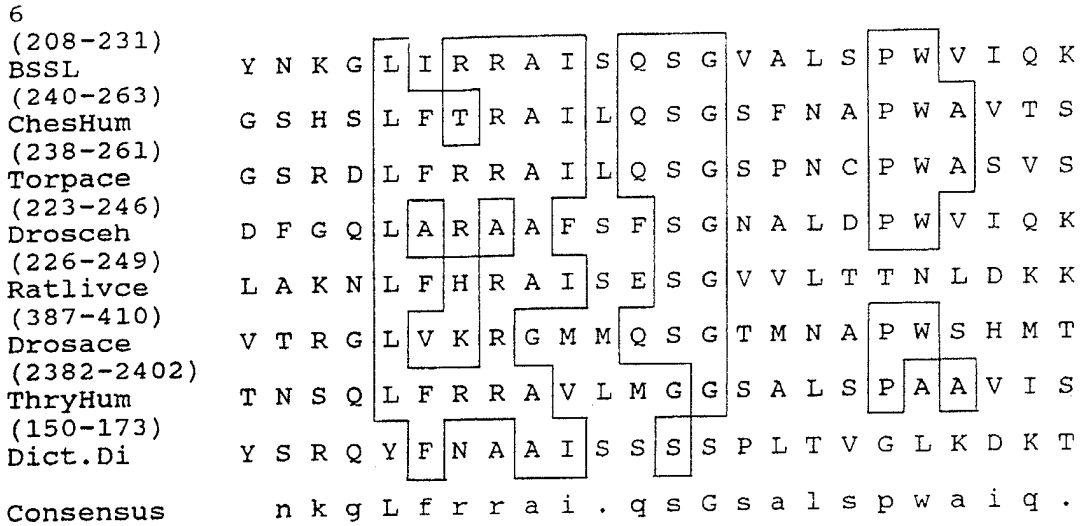
6
(208-231) BSSL Y N K G L I R R A I S Q S G V A L S P W V I Q K
(240-263) ChesHum G S H S L F T R A I L Q S G S F N A P W A V T S
(238-261) Torpace G S R D L F R R A I L Q S G S P N C P W A S V S
(223-246) Drosceh D F G Q L A R A A F S F S G N A L D P W V I Q K
(226-249) Ratlivce L A K N L F H R A I S E S G V V L T T N L D K K
(387-410) Drosace V T R G L V K R G M M Q S G T M N A P W S H M T
(2382-2402) ThryHum T N S Q L F R R A V L M G G S A L S P A A V I S
(150-173) Dict.Di Y S R Q Y F N A A I S S S S P L T V G L K D K T
Consensus n k g L f r r a i . q s G s a l s p w a i q .
7
(285-297) BSSL V G F V P V I D G D F I P
(316-328) ChesHum V N F G P T V D G D F L T
(314-326) Torpace F S F V P V I D G E F F P
(298-310) Drosceh A P F S P V L E P S D A P
(298-306) Ratlivce . . . . T V I D G V V L P
(466-477) Drosace P S . A P T I D G A F L P
(2458-2468) ThryHum . . W G P V I D G H F L R
(224-236) Dict.Di T I W S P V I D G D A F I
Consensus . . f . p v . d g d f . p
1
5aa
2
22aa
3
4
5
6
54aa
7

Fig. 7.


10
30
50
AKLGAVYTEGGFVEGVNKKLGLLGDSVDIFKGIPFAAPTKALENPQPHPGWQGTLKAKNF
70
90
110
KKRCLQATITQDSTYGDEDCLYLNIWVPQGRKQVSRDLPVMIWIYGGAFLMGSGHGANFL
130
150
170
NNYLYDGEEIATRGNVIVVTFNYRVGPLGFLSTGDANLPGNYGLRDQHMAIAWVKRNIAA
190
210
230
FGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSGVALSPWVIQKNPLFWAKKV
250
270
290
AEKVGCPVGDAARMAQCLKVTDPRALTLAYKVPLAGLEYPMLHYVGFVPVIDGDFIPADP
exon a
310
330
350
INLYANAADIDYIAGTNNMDGHIFASIDMPAINKGNKKVTEEDFYKLVSEFTITKGLRGA
370
390
410
KTTFDVYTESWAQDPSQENKKKTVVDFETDVLFLVPTEIALAQHRANAKSAKTYAYLFSH
exon-b
430
450
470
PSRMPVYPKWVGADHADDIQYVFGKPFATPTGYRPQDRTVSKAMIAYWTNFAKTGDPNMG
exon c
490
510
530
DSAGPTHWEPYTTENSGYLEITKKMGSSSMKRSLRTNFLRYWTLTYLALPTVTDQEATPV
exon d
#
550
570
590
PPTGDSEATPVPPTGDSETAPVPPTGDSGAPPVPPTGDSGAPPVPPTGDSGAPPVPPTGD
610
630
650
SGAPPVPPTGDSGAPPVPPTGDSGAPPVPPTGDSGAPPVPPTGDAGPPPVPPTGDSGAPP
670
690
710
VPPTGDSGAPPVTPTGDSETAPVPPTGDSGAPPVPPTGDSEAAPVPPTDDSKEAQMPAVI
RF

INTERNATIONAL SEARCH REPORT

International Application No PCT/SE 91/00381

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC5: C 07 K 13/00, C 12 N 9/20, C 12 N 9/16, C 12 N 15/55
A 61 K 37/54

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
IPC5	C 07 K; C 12 N; A 61 K

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in Fields Searched[8]

SE,DK,FI,NO classes as above

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category *	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
P,X	Dialog Information Services, File 155: Medline 67-91, NLM accession number 91006144, Nilsson J et al: "cDNA cloning of human-milk bile-salt-stimulated lipase and evidence for its identity to pancreatic carboxylic ester hydrolase" & Eur J Biochem (GERMANY) Sep 11 1990, 192 (2) p543-50 --	1-30
A	WO, A1, 8500381 (CELLTECH LIMITED) 31 January 1985, see the whole document --	1-30

---

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
29th August 1991	1991 -09- 09
International Searching Authority	Signature of Authorized Officer
SWEDISH PATENT OFFICE	Mikael Bergstrand

Form PCT/ISA/210 (second sheet) (January 1985)

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO.PCT/SE 91/00381

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the Swedish Patent Office EDP file on 91-06-27
The Swedish Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A1- 8500381	85-01-31	AU-D-	3107684	85-02-07
		EP-A-	0131418	85-01-16
		GB-A-	2142337	85-01-16